



The Open Biomedical Engineering Journal

Content list available at: <https://openbiomedicalengineeringjournal.com>



RESEARCH ARTICLE

Multi-class Classification of Gastrointestinal Diseases using Deep Learning Techniques

X. Anitha Mary^{1,*}, A. Peniel Winifred Raj¹, C. Suganthi Evangeline², T. Mary Neebha², Vinoth Babu Kumaravelu³ and P. Manimegalai⁴

¹Department of Robotics Engineering, Karunya Institute of Technology and Sciences, Coimbatore, Tamil Nadu, India

²Department of Electronics and Communication Engineering, Karunya Institute of Technology and Sciences, Coimbatore, Tamil Nadu, India

³Department of Communication Engineering, School of Electronics Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

⁴Department of Biomedical Engineering, Karunya Institute of Technology and Sciences, Coimbatore, Tamil Nadu, India

Abstract:

Objective:

Gastrointestinal problems have claimed the lives of about two million individuals worldwide. Video endoscopy is one of the most important tools for detecting gastrointestinal diseases (GD). However, it has the disadvantage of producing a huge quantity of images that should be analysed by medical professionals, which demands more time. This makes manual diagnosis challenging, encouraging research on computer-assisted methods for accurately identifying all of the produced images in a short amount of time. The goal of this research is to develop an artificial intelligence technique to classify different gastrointestinal problems.

Materials and Methods:

In developing a framework for classifying gastrointestinal diseases, the proposed methodology is unique. VGG-19 and ResNet-50, two deep learning networks, were used to evaluate their ability to recognise a dataset of lower GD. To validate the proposed models, a Kvasir dataset of 3500 images was employed, comprising 500 images from each of the seven classes.

Results:

ResNet-50 had the greatest validation accuracy of 96.81%, recall of 95.28%, precision of 95%, and 94.85% of F1 score, whereas VGG-19 had the highest validation accuracy of 94.21%, recall of 94%, precision of 94.28%, and 94% of F1 score.

Conclusion:

The Convolutional Neural Network (CNN) was used to extract shape, colour, and texture features. Softmax assists in the classification of seven types of GD by passing these features into fully connected layers. Two models performed equally well and yielded promising results.

Keywords: Classification, Deep learning, Gastrointestinal diseases, Kvasir dataset, ResNet-50, VGG-19.

Article History

Received: May 18, 2022

Revised: December 09, 2022

Accepted: December 21, 2022

1. INTRODUCTION

In comparison to other malignancies, gastrointestinal cancer is the most prevalent disease. A total of 1.8 million people die each year from gastrointestinal disorders, according to the World Health Organization (WHO) [1]. Gastroscopy tools are not appropriate for detecting and examining GD infections, such as polyps, ulcers, and bleeding, because of

their intricacy. Wireless capsule endoscopy was developed in 2000 to tackle the difficulty with gastroscopy equipment [2]. According to the 2018 annual report, the above equipment was used on almost one million patients [3].

GD infections, particularly in the small intestine, can be identified, but the overall method takes a long time to evaluate the video frames obtained from each patient. On the other hand, physicians face difficulties that need an excess of time to evaluate all of the data [4]. In various medical fields, artificial intelligence (AI) technologies have shown great potential in aiding humans in visualising disorders that are not evident to

* Address correspondence to this author at the Department of Robotics Engineering, Karunya Institute of Technology and Sciences, Coimbatore, Tamil Nadu, India; E-mail: anithamary@karunya.edu

the naked eye. The application of two CNN models for the identification and classification of gastrointestinal disorders is the subject of this article. Machine learning is highly useful in medical diagnoses, including lung disease, chest, etc. [20 - 22].

By incorporating engineering into learning, CNN has demonstrated its remarkable ability to abstract features. In medical image diagnostics, deep learning algorithms have proved that they can outperform specialists [5]. Consequently, computer-assisted diagnoses for endoscopic images based on deep learning algorithms are projected to attain diagnostic accuracy comparable to that of qualified specialists [6]. Igarashi *et al.* presented the AlexNet model for classifying upper gastrointestinal organs, which had a 96.5% accuracy [7].

Biniaz *et al.* employed factorization analysis, a technique for speeding up endoscopic screening evaluation [8]. The sliding window mechanism with singular value decomposition was also utilized, which offered an overall accuracy of 92%. Cogan *et al.* employed a variety of CNN pre-trained models, including Inception V2, Inception V4, NasNet, and ResNet-V2 [9]. With Inception V4, a classification accuracy of 93.8% can be achieved. AbraAyidzoe *et al.* suggested a Gabor capsule network that classifies the Kvasir dataset images with a 93.65% accuracy [10]. Song *et al.* employed CNN to classify polyps [11]. Mohaputra *et al.* employed the Wavelet transform in conjunction with the CNN model to reach a 93.65% accuracy [12]. To extract information from the image caption and also in the text, a sequential approach and a multimodal one were proposed by Aggarwal *et al.* [13] (Table 1).

The rest of the article is organized as follows: the materials and methodologies are discussed in Section 2. The simulation results are discussed in Section 3. The conclusions are

highlighted finally.

2. MATERIALS AND METHODOLOGIES

Computer-assisted automated detection of GD is becoming a hot topic of research. In this section, VGG- 19 and ResNet-50 are described in detail. Fig. (1) depicts a generic block diagram for early diagnosis of GD. To extract features from the image, CNN is employed. Diagnosis and classification are aided by the fully connected layers [20 - 22].

2.1. VGG-19 Model

VGG-19 is a CNN with 19 layers trained on over a million images in the ImageNet database. This model has six core structures, each of which is composed of various connected convolutional layers and fully-connected layers. Fig. (2) shows the architecture of VGG-19, which is employed in this investigation. A fixed image size of 224 x 224 RGB is used in the architecture. It makes use of a 3 x 3 kernel. Spatial paddings are employed before 2 x 2 max pooling to preserve the spatial resolution. Rectified linear unit (ReLU) is used to increase classification while reducing computing time. This model is employed as a pre-processing model, and it has better network depth than the classic CNN. Multiple convolutional layers and activation layers with non-linear functionality provide a superior and alternating structure. The layer structure may abstract image features, then downsample them using max-pooling and alter the ReLU to indicate the major value in the image region as the area's pooled value. The downsampling layer increases the anti-distortion capability of the network. By considering only the important features, it reduces the parameters. An Adam optimizer is also utilised in this study, with a learning rate of 0.0001.

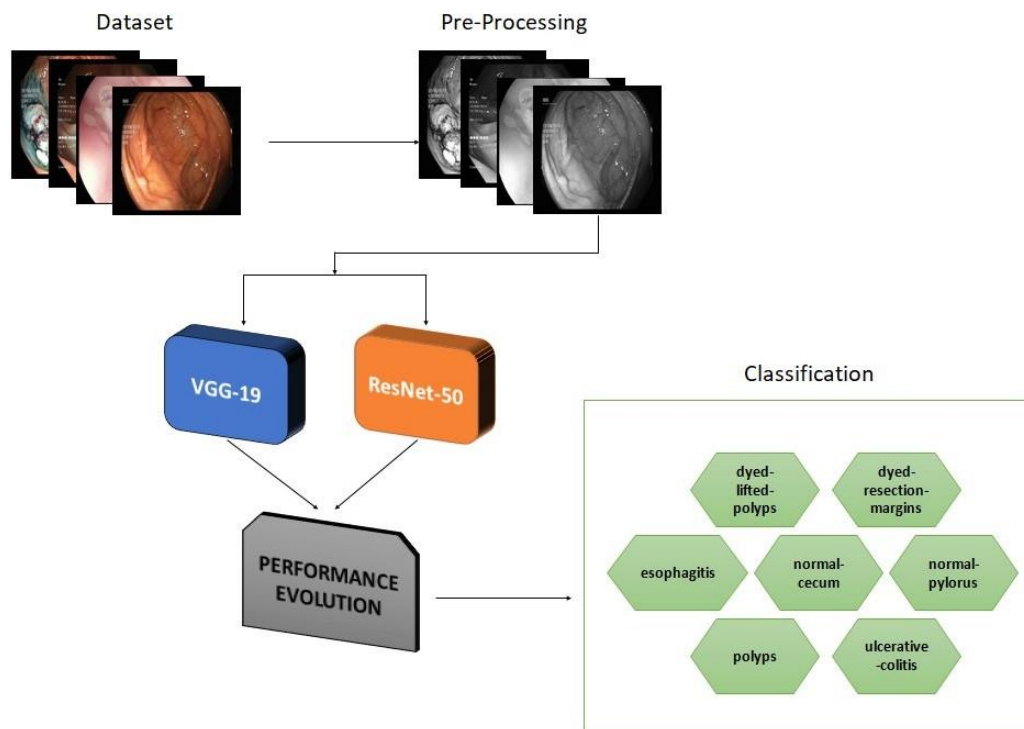


Fig. (1). General block diagram for early diagnosis of GD.

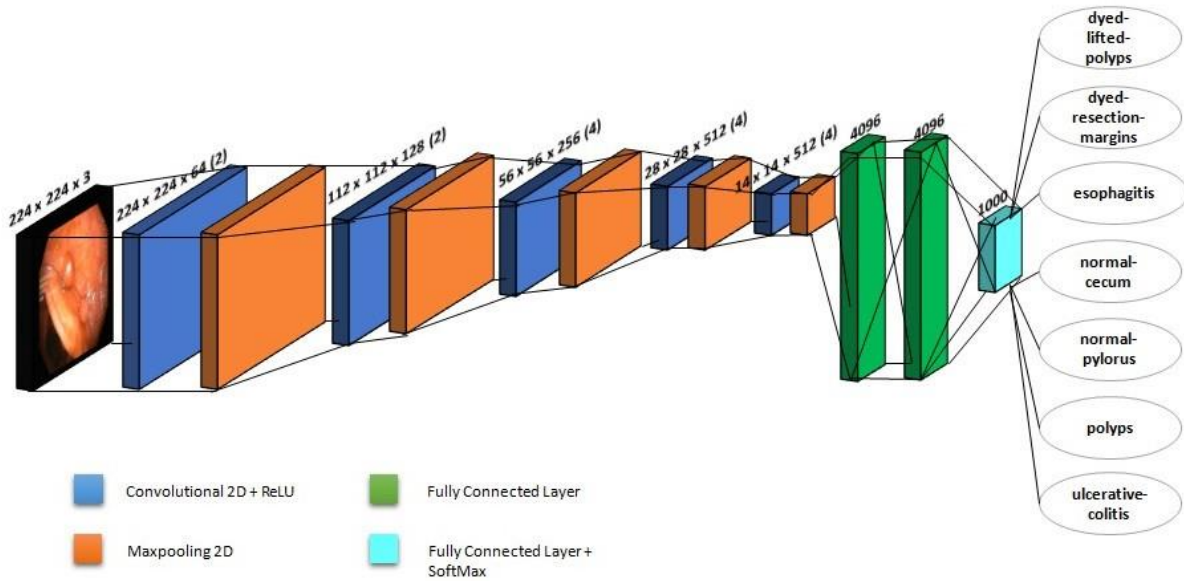


Fig. (2). Architecture of VGG-19 model.

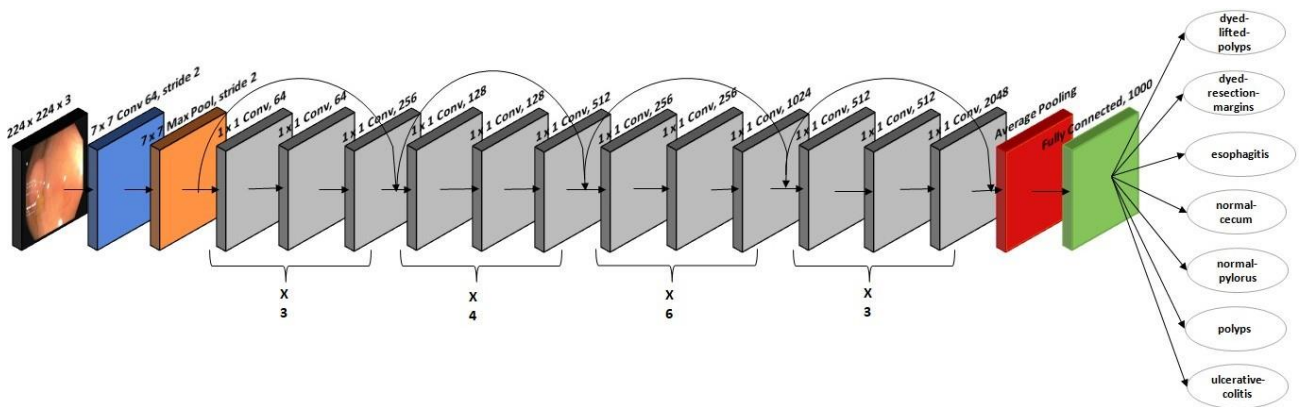


Fig. (3). Architecture of ResNet-50 model.

2.2. ResNet-50 Architecture

A residual CNN model with 177 layers is used in ResNet-50 architecture. ResNet-50 won the image classification competition in 2015. The ResNet-50 network is mainly used in computer vision applications. The architecture of ResNet-50 is divided into four stages. The network can accept an image with a height and width that are multiples of 32 and a channel width of 3; 224 x 224 x 3 is the input size specified in the design. Every ResNet architecture adopts a kernel size of 7 x 7 for initial convolution and 3 x 3 for max pooling. Stage 1 contains three residual blocks, each with three levels. The kernels used to execute convolution operations in all three levels of stage 1's block are 64, 64, and 128. The channel width doubles, and the input size becomes half as we progress from one stage to the next. The network contains an

average pooling layer, especially towards the end, succeeded by a fully connected layer of 1000 neurons. Fig. (3) shows the ResNet-50's architecture, which is used to classify 3500 images into seven lower digestive tract disorders. It includes 16 input layers and 177 layers between them. After that, only positive data is then transmitted *via* the convolutional layers. In this study, an stochastic gradient descent (SGD) optimizer with a learning rate of 0.00001 is employed.

2.3. Dataset under Study

The Kvasir dataset, which comprises images of the gastrointestinal system acquired with endoscopic equipment, was used in this investigation. Fig. (4) shows sample images with subclasses from the proposed study, which includes seven classes with 500 images each. The images are resized to 224 x 224 pixels.

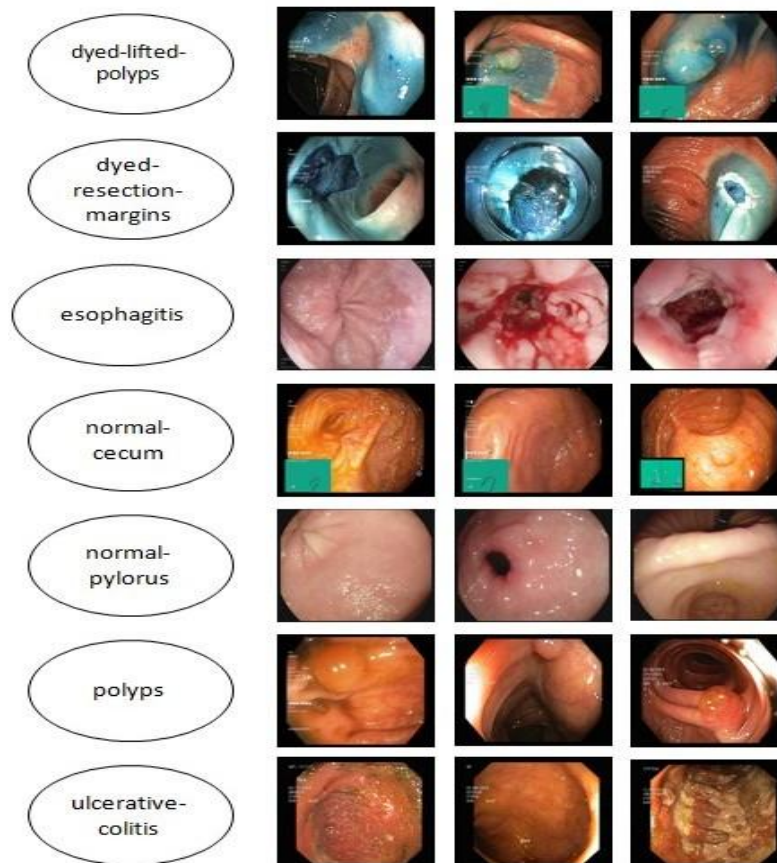


Fig. (4). Sample images for various gastrointestinal tract disease.

2.4. Preprocessing and Augmentation Techniques

Due to the increased complexity of feature extraction, noise and artifacts caused by light reflections, camera angles, and the mucosal membranes covering internal organs degrade CNN performance. Therefore, researchers have been interested in optimization approaches to enhance image quality. After resizing the image to 224×224 pixels, an average filter is used as the final step in the improvement process. It replaces each pixel by averaging it with its neighbours, a procedure that is repeated for all the pixels in the image.

2.5. Convolution Layers

Numerous features, including shape, texture, and colour, are present in the gastrointestinal dataset. When extracting images from a video, many images do not include the disease, and the disease features only exist in a few images that may be overlooked by the radiologist and the specialists. Hence, manual extraction of features requires significant skill. However, in the case of CNN algorithms, convolutional layers are used to extract representative aspects of each disease. VGG-19 is a CNN with 19 layers, and the ResNet50 CNN model has 177 layers. The classification layers have the 9216 representative features that the convolutional layers extract from each image and represent as feature maps.

2.6. Normalization and Dropout Technology

In order to speed up training procedures for deep neural networks, normalisation is one of the approaches utilised during training. Gradient descent converges normalisation and assists in selecting the proper learning rate. The learning rate is more challenging and requires more time without the normalisation process. In our research, the mean of the entire training set for each pixel was subtracted during the image normalisation process. Data was centered, and each feature's variance was set to one by computing the dataset's variance and dividing it by each pixel. Millions of parameters are generated by CNNs, which causes overfitting. In order to lessen overfitting, CNN uses a dropout method. In our research, the dropout approach was used to stop 50% of the neurons in ResNet-50 and 40% for VGG 19 in each iteration. As a result, each iteration of the network used a distinct set of parameters. Hence, the training time doubled with the dropout technique.

2.7. Transfer Learning and Optimizer

Transfer learning is based on using a dataset to train on a particular problem and then applying that knowledge to another problem with a different dataset. In order to apply what has been learned to a different task, transfer learning selects the pre-trained model and the size of the challenge. Overfitting is also avoided through transfer learning. In this study, the weights of ResNet-50 and VGG 19 were adjusted using

transfer learning. The gastrointestinal dataset was used to apply the learning from the ImageNet dataset to the ResNet-50 and VGG 19 models. Optimizers were primarily utilized to tune up the weights, bias, and learning rate of a model. The Adam optimizer was used by AlexNet, whereas the (SGD) optimizer was used by ResNet- 50.

3. RESULTS AND DISCUSSION

In cross-validation, 80% of the dataset was used for training, and the remaining 20% was used for testing. A coLab Pro environment is used to simulate the whole model. The confusion matrix produced from ResNet-50 and VGG-19 is shown in Figs. (5 and 6).

ResNet-50 showed the greatest validation accuracy of 96.81%, whereas VGG-19 had the highest validation accuracy of 94.21%. Fig. (7) depicts the models' overall metrics, whereas Fig. (8) depicts the metrics derived for each disease.

The accuracy and loss plots for VGG 19 and ResNet 50 are shown in (Figs. 9 and 10).

The proposed models were compared to models already published in the literature, as shown in Table 1. ResNet-50 was reported to have a validation accuracy of 99.82%, which was higher than the other models. Fig. (9) displays a graphical representation of performance metrics of models' understudy compared to other models.

	dyed-lifted-polyps	dyed-resection-margins	esophagitis	normal-cecum	normal-pylorus	polyps	ulcerative-colitis
dyed-lifted-polyps	96	3	0	0	0	1	0
dyed-resection-margins	14	86	0	0	0	0	0
esophagitis	0	0	100	0	0	0	0
normal-cecum	0	0	0	97	0	3	0
normal-pylorus	0	0	0	0	97	2	1
polyps	0	0	1	2	0	97	0
ulcerative-colitis	1	0	1	3	0	3	93

Fig. (5). Confusion matrix resulted from ResNet-50.

	dyed-lifted-polyps	dyed-resection-margins	esophagitis	normal-cecum	normal-pylorus	polyps	ulcerative-colitis
dyed-lifted-polyps	91	9	0	0	0	0	0
dyed-resection-margins	8	91	0	0	0	1	0
esophagitis	0	0	99	0	1	0	0
normal-cecum	0	0	0	91	0	9	0
normal-pylorus	0	0	0	0	99	1	0
polyps	0	0	0	2	2	96	0
ulcerative-colitis	1	0	1	0	0	7	91

Fig. (6). Confusion matrix resulted from VGG-19.

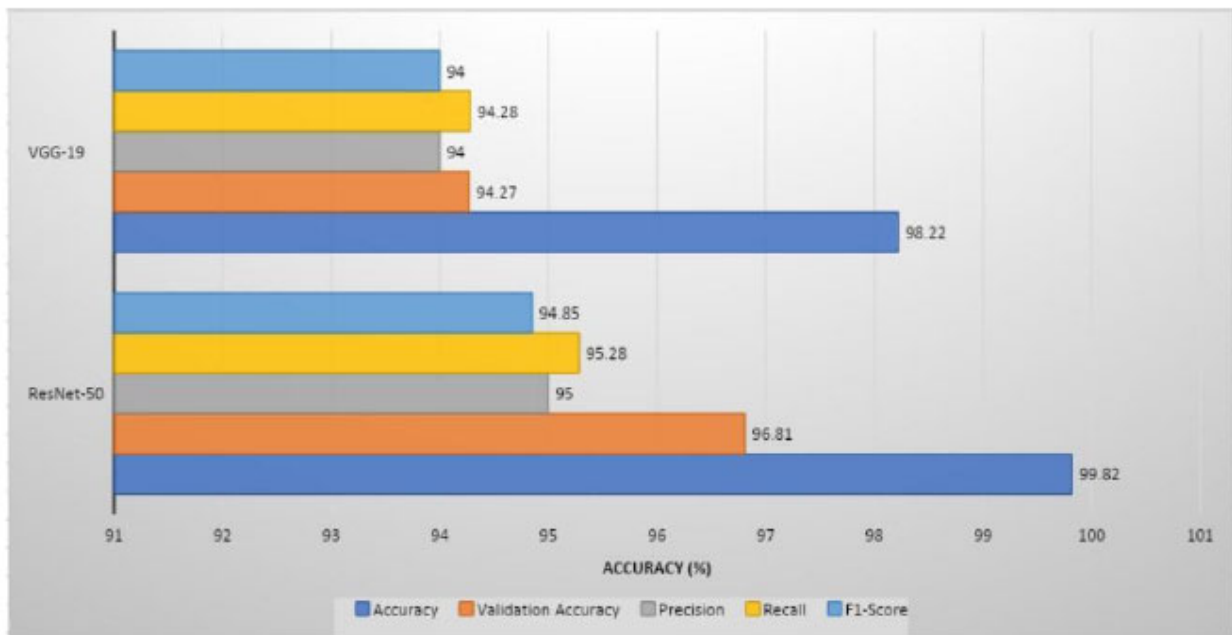


Fig. (7). Classification performance metrics for two models.

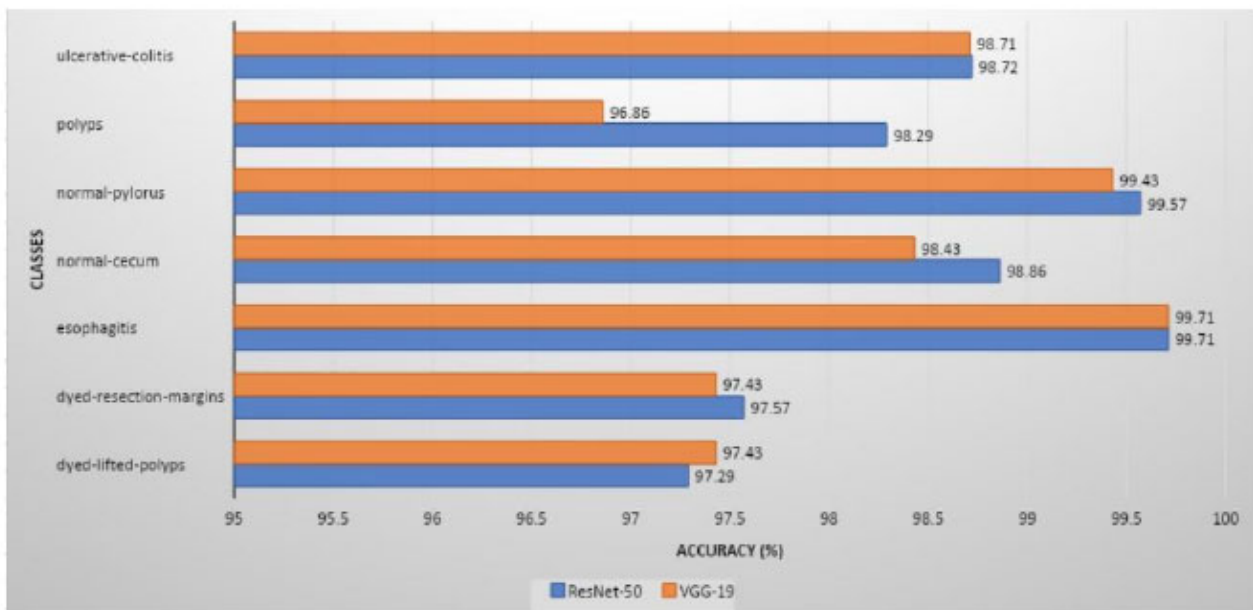


Fig. (8). Classification accuracy achieved by the proposed schemes for each disease.

Table 1. Comparative analysis of previous works with models' understudy.

Related Studies	Accuracy (%)	Sensitivity (%)	Specificity (%)
Godkhindi and Gowda [14]	88.56	88.77	87.35
Pozdeev et al. [15]	88	93	82
Bour et al. [16]	87.1	87.1	93
Min et al. [17]	78.4	83.3	70.1

Related Studies	Accuracy (%)	Sensitivity (%)	Specificity (%)
Fonoll'a <i>et al.</i> [18]	90.2	90.1	90.3
Mosleh <i>et al.</i> – AlexNet [19]	97	96.8	99.2
Mosleh <i>et al.</i> – ResNet-50 [19]	95	94.8	98.8
ResNet-50 (Model understudy)	99.82	95.23	99.16
VGG-19 (Model understudy)	98.22	94.3	99

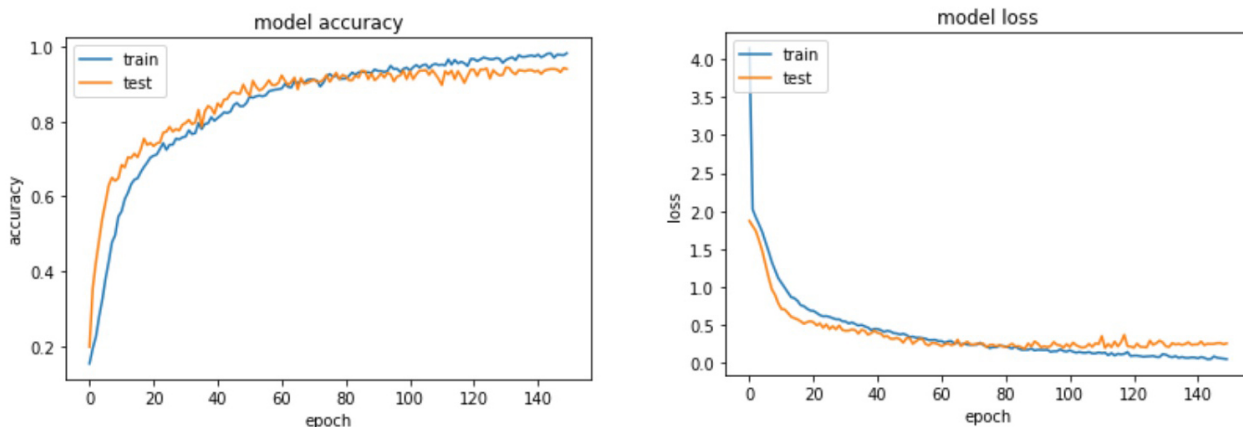


Fig. (9). Accuracy and loss plots for VGG 19 model.

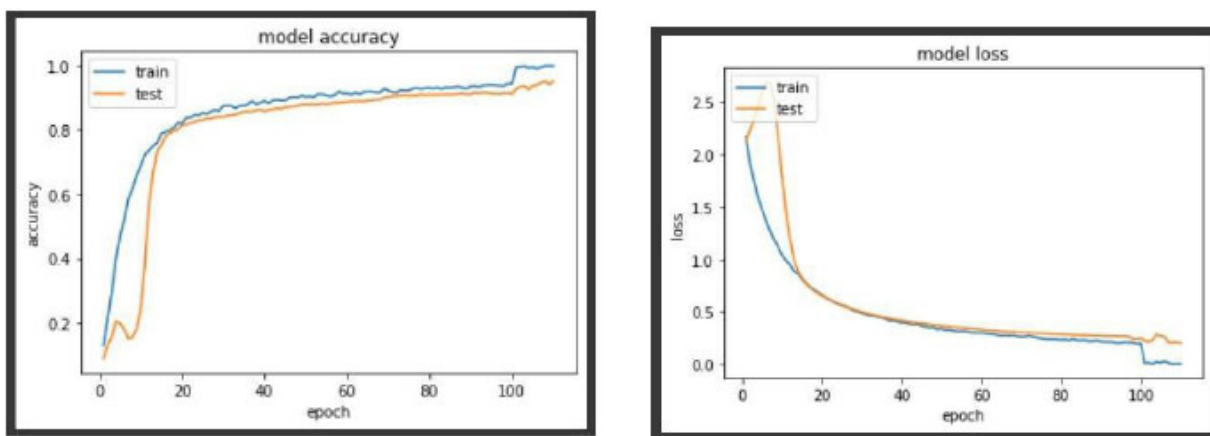


Fig. (10). Accuracy and loss plots for ResNet 50 model.

CONCLUSION

Using the Kvasir dataset, this research proposes a robust method for the classification of gastrointestinal disorders. Deep learning algorithms can help to lessen the chance of emerging cancer by assisting in early detection and reducing the need for unnecessary begin tumour excision. Despite the fact that video endoscopy is the most prevalent investigative approach for finding gastrointestinal polyps, a number of human aspects contribute to inaccurate gastrointestinal illness diagnosis. VGG-19 and ResNet-50, two deep-learning models studied in this study, can direct doctors' attention to the vital areas that

may have been neglected previously. A total of 9216 characteristics were retrieved and sent to the 1000 neurons in fully connected layers. Each image was categorised into one of the seven categories of gastrointestinal disorders by the softmax layer, which generated seven classes. The VGG 19 and ResNet50 models produced results that were quite promising. Therefore, future applications of advanced deep learning algorithms are suggested. However, the algorithm's weakest element is the need for manual data marking. As accurate disease diagnosis is frequently challenging for even humans, the network may therefore inherit certain errors from an analyst. One method to overcome this restriction is to use a

larger dataset that has been labelled by a broader group of experts.

LIST OF ABBREVIATIONS

GD	=	Gastrointestinal Diseases
CNN	=	Convolutional Neural Network
WHO	=	World Health Organization

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used in the studies that are the basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The data supporting the findings of the article is available in the Kaggle at <https://www.kaggle.com/datasets/plhalvorsen/kvasir-v2-a-gastrointestinal-tract-dataset>.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflicts of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

- [1] "Latest global cancer data", Available from: https://www.iaarc.fr/wpcontent/uploads/2018/09/pr263_E.pdf
- [2] T. Rahim, M.A. Usman, and S.Y. Shin, "A survey on contemporary computer-aided tumor, polyp, and ulcer detection methods in wireless capsule endoscopy imaging", *Comput. Med. Imaging Graph.*, vol. 85, p. 101767, 2020. [http://dx.doi.org/10.1016/j.compmedimag.2020.101767] [PMID: 32966967]
- [3] A. Liaqat, M.A. Khan, J.H. Shah, M. Sharif, M. Yasmin, and S.L. Fernandes, "Automated ulcer and bleeding classification from WCE images using multiple features fusion and selection", *J. Mech. Med. Biol.*, vol. 18, no. 4, p. 1850038, 2018. [http://dx.doi.org/10.1142/S0219519418500380]
- [4] M. Islam, B. Chen, J.M. Spraggins, R.T. Kelly, and K.S. Lau, "Use of single-cell - omics technologies to study the gastrointestinal tract and diseases, from single-cell identities to patient features", *Gastroenterology*, vol. 159, no. 2, pp. 453-466.e1, 2020. [http://dx.doi.org/10.1053/j.gastro.2020.04.073] [PMID: 32417404]
- [5] T.N. Sainath, B. Kingsbury, G. Saon, H. Soltan, A. Mohamed, G. Dahl, and B. Ramabhadran, "Deep convolutional neural networks for large-scale speech tasks", *Neural Netw.*, vol. 64, pp. 39-48, 2015. [http://dx.doi.org/10.1016/j.neunet.2014.08.005] [PMID: 25439765]
- [6] V. Mnih, K. Kavukcuoglu, D. Silver, A.A. Rusu, J. Veness, M.G. Bellemare, A. Graves, M. Riedmiller, A.K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning", *Nature*, vol. 518, no. 7540, pp. 529-533, 2015. [http://dx.doi.org/10.1038/nature14236] [PMID: 25719670]
- [7] S. Igarashi, Y. Sasaki, T. Mikami, H. Sakuraba, and S. Fukuda, "Anatomical classification of upper gastrointestinal organs under various image capture conditions using AlexNet", *Comput. Biol. Med.*, vol. 124, p. 103950, 2020. [http://dx.doi.org/10.1016/j.compbmed.2020.103950] [PMID: 32798923]
- [8] A. Biniiaz, R.A. Zoroofi, and M.R. Sohrabi, "Automatic reduction of wireless capsule endoscopy reviewing time based on factorization analysis", *Biomed. Signal Process. Control*, vol. 59, p. 101897, 2020. [http://dx.doi.org/10.1016/j.bspc.2020.101897]
- [9] T. Cogan, M. Cogan, and L. Tamil, "MAPGI: Accurate identification of anatomical landmarks and diseased tissue in gastrointestinal tract using deep learning", *Comput. Biol. Med.*, vol. 111, p. 103351, 2019. [http://dx.doi.org/10.1016/j.compbmed.2019.103351] [PMID: 31325742]
- [10] A.A. Mighty, and Y. Yongbin, "Gabor capsule network with preprocessing blocks for the recognition of complex images", *Mach. Vis. Appl.*, vol. 32, p. 91, 2021. [http://dx.doi.org/10.1007/s00138-021-01221-6]
- [11] E.M. Song, B. Park, C.A. Ha, S.W. Hwang, S.H. Park, D.H. Yang, B.D. Ye, S.J. Myung, S.K. Yang, N. Kim, and J.S. Byeon, "Endoscopic diagnosis and treatment planning for colorectal polyps using a deep-learning model", *Sci. Rep.*, vol. 10, no. 1, pp. 30-10, 2020. [http://dx.doi.org/10.1038/s41598-019-56697-0] [PMID: 31913337]
- [12] S. Mohapatra, J. Nayak, M. Mishra, G.K. Pati, B. Naik, and T. Swarnkar, "Wavelet transform and deep convolutional neural network-based smart healthcare system for gastrointestinal disease detection", *Interdiscip. Sci.*, vol. 13, no. 2, pp. 212-228, 2021. [http://dx.doi.org/10.1007/s12539-021-00417-8] [PMID: 33566337]
- [13] A. Aggarwal, V. Sharma, A. Trivedi, M. Yadav, C. Agrawal, D. Singh, V. Mishra, and H. Gritli, "Two-way feature extraction using sequential and multimodal approach for hateful meme classification", *Complexity*, vol. 2021, pp. 1-7, 2021. [http://dx.doi.org/10.1155/2021/5510253]
- [14] A.M. Godkhindi, and R.M. Gowda, "Automated detection of polyps in CT colonography images using deep learning algorithms in colon cancer diagnosis", *Proceedings of the 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, 2017pp. 1722-1728 [http://dx.doi.org/10.1109/ICECDS.2017.8389744]
- [15] A.A. Pozdeev, N.A. Obukhova, and A.A. Motyko, "Automatic analysis of endoscopic images for polyps detection and segmentation", *Proceedings of the 2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIcon Rus)*, 2019pp. 1216-1220. [http://dx.doi.org/10.1109/EIconRus.2019.8657018]
- [16] A. Bour, C. Castillo-Olea, B. Garcia-Zapirain, and S. Zahia, "Automatic colon polyp classification using convolutional neural network: A case study at Basque country", *Proceedings of the 2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, 2019pp. 1-5. [http://dx.doi.org/10.1109/ISSPIT47144.2019.9001816]
- [17] M. Min, S. Su, W. He, Y. Bi, Z. Ma, and Y. Liu, "Computer-aided diagnosis of colorectal polyps using linked color imaging colonoscopy to predict histology", *Sci. Rep.*, vol. 9, no. 1, pp. 2881-2888, 2019. [http://dx.doi.org/10.1038/s41598-019-39416-7] [PMID: 30814661]
- [18] R. Fonollá, F. Van Der Sommen, R.M. Schreuder, E.J. Schoon, and P.H. De With, "Multi-modal classification of polyp malignancy using CNN features with balanced class augmentation", *Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, 2019pp. 74-78. [http://dx.doi.org/10.1109/ISBI.2019.8759320]
- [19] M.H. Al-Adhaileh, M.S. Ebrahim, W.A. Fawaz, H.H.A. Theyazn, N. Alsharif, A.A. Ahmed, and U. M.Irfan, "Deep learning algorithms for detection and classification of gastrointestinal diseases", *Complexity*, vol. 2021, p. 6170416, 2021. [http://dx.doi.org/10.1155/2021/6170416]
- [20] L. Devnath, P. Summons, S. Luo, D. Wang, K. Shaikat, I.A. Hameed, and H. Aljuaid, "Computer-aided diagnosis of coal workers' pneumoconiosis in chest x-ray radiographs using machine learning: A systematic literature review", *Int. J. Environ. Res. Public Health*, vol. 19, no. 11, p. 6439, 2022. [http://dx.doi.org/10.3390/ijerph19116439] [PMID: 35682023]

[21] D. Wang, "Automated pneumoconiosis detection on chest x-rays using cascaded learning with real and synthetic radiographs", *Digital Image Computing: Techniques and Applications (DICTA)*, vol. 2020, pp. 1-6, 2020.

[22] [http://dx.doi.org/10.1109/DICTA51227.2020.9363416]
D. Liton, P.S. Suhuailuo, and D.T. Wang, "Classification in chest radiographs using deep convolutional neural networks", *Int. J. Adv. Sci. Eng. Tech.*, vol. 6, no. 1, pp. 2321-9009, 2018. Available from: <http://iraj.in>

© 2023 Mary *et al.*

This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Public License (CC-BY 4.0), a copy of which is available at: <https://creativecommons.org/licenses/by/4.0/legalcode>. This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.